Semaine d'Etude Mathématiques et Entreprises 7 : Détection d'ilôtage dans un réseau électrique

Zoe Amblard, Aalae Benki, Camilo A. Garcia Trillos, Thi Nguyet Nga Ta, Thi Thanh Phuong Truong

7ème Semaine d'Étude Maths-Entreprises
4-8 Novembre 2013, XLIM (Limoges)

Détection d'ilôtage dans un réseau électrique

Zoe Amblard[a]
Aalae Benki[c]
Camilo Andrés Garcia Trillos[e]
Thi Nguyet Nga Ta [b]
Thi Thanh Phuong Truong [d]

[a] *XLIM - UMR CNRS n7252 - Université de Limoges, 87000 Limoges, France*
[b] *XLIM - UMR CNRS n7252 - Université de Limoges, 87000 Limoges, France*
[c] *Inria - Université de Nice Sophia Antipolis, 06902 Sophia Antipolis, France*
[d] *Laboratoire de Mathématiques, Université d'Avignon, 84000 Avignon, France*
[e] *Laboratoire J.A. Dieudonné, Université Nice Sophia Antipolis, 06108 Nice, France*

Sujet proposé par:

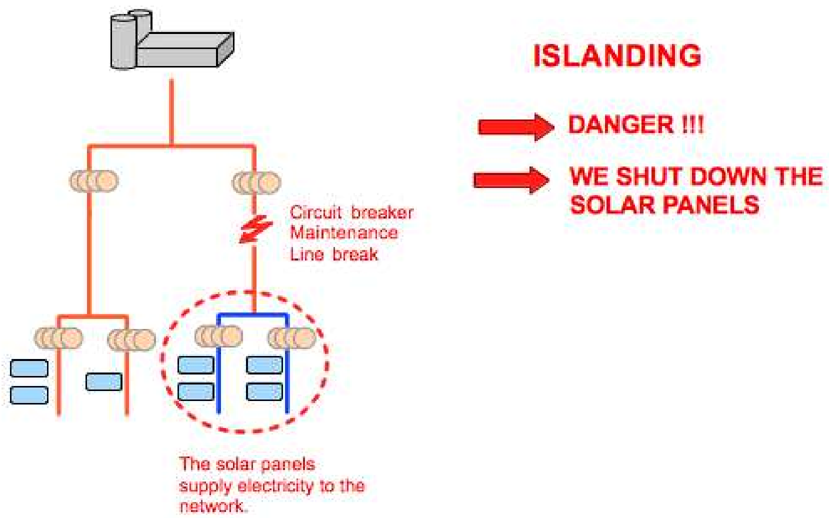
Correspondant:

Florent CADOUX (Chaire ERDF, Fondation Grenoble INP)

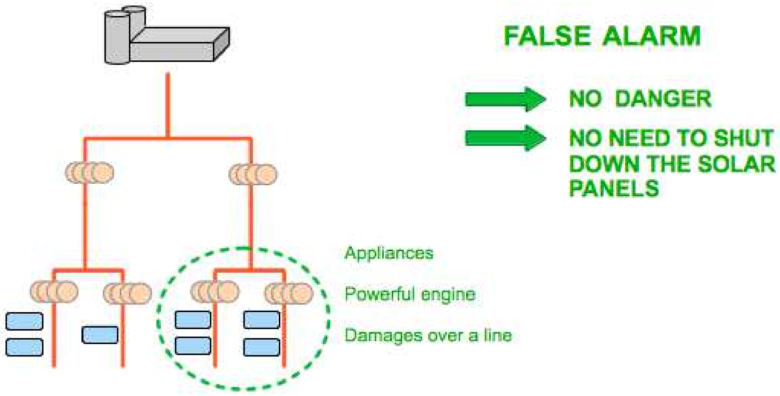
1 Introduction

In recent years, the significant increase in number of connections to photovoltaic installations in the public networks has caused greater risks. Methods involving in detecting situations such as islanding, damage on the transport network, and normal situations without special event arise have been exploited, however, each method has its disavantages. In this work, by using the statistical and signal processing tools, we can consider and appreciate the change of data of the simulated model. Then we apply the CART algorithm to classify the above phenomena.

What is islanding ?
When an incident occurs in the distribution network, a part of this network can be (willingly or not) isolated. If it also contains devices to generate energy such as solar panels, then this area is still independently supplied in electricity. This situation is called islanding. The distribution network is no longer able to regulate the current through the island. Such a situation may be dangerous, and it is best to stop the independent power generators to prevent unregulated production of electricity. The detection of islanding condition is based on a sudden change in the network load. However, in real situations, it is often very difficult to distinguish between a false alarm triggered by a disturbance on the transmission (line break, neighboring engine, noise, etc. ..), which is called a false positive, and islanding.



alarm.jpg



There are three type of detection methods that could allow us to recognise islanding from false positives :

- Passive methods detect transient changes on parameters such as voltage, frequency or harmonic distortion. These modifications on the grid are usually extremely brief and they can also be caused by a false positive such as the start of a neighboring engine. Consequently, the most difficult here is to find proper differenciation criterions.
- Active methods consist in injecting deliberately a signal into the line and detect modifications in order to determinate if the local circuit is still connected to the main grid.
- Communication methods require specific devices to allow the solar panels to exchange informations with the main grid or between each other.

Our method belongs to the passive methods category.

2 Procedure to define a passive detection method

We propose a statistical approach to define an islanding detection criterion. In essence, we reformulate the question as a classification problem of a set of indicator functions (that we will be detailed in the following) into two classes characterized by whether there is presence or not of an islanding condition. Here, the term indicator is used to denote a set of functions transforming a continuous voltage (or current) signal into real values. Evidently the criterion definition is given in terms of these indicator functions, and hence its final implementability is directly associated to the used indicators.

Let us first introduce the notation we use in the following:

- T: Time window of observation.
- M: Number of distinct observations.
- S_i, $i = 1, \ldots, M$: continuous signals of time duration T. Without loss of generality, we consider $S_i : [0, T] \to \mathbb{R}$.
- N: Number of distinct indicators.
- f^j, $j = 1, \ldots, N$: indicators functions with $f_j : [0, T] \to \mathbb{R} \to \mathbb{R}$.
- X_i^j, $k = 1, \ldots, M \cdot N$: shorthand for the indicators applied to the available signals, i.e. $X_i^j := f^j(V_i)$. They constitute a set of M observations of N explanation variables.
- Y_i: binary variable denoting the presence (1) or absence (0) of the islanding condition in the i-th observation.

The procedure we propose is composed of the following basic steps

- Database construction
- Definition of indicator functions
- Classification

2.1 Database construction

The main input behind any statistical or learning method is information. In general, the quality of the final classification is limited by the quality of the initial database.

Unfortunately, there is no database registering islanding instances, mostly because, historically, it has been a very rare phenomenon. As the number of distributed generators in the distribution networks increase, so will the probability of arrival of islanding. For now, we have to look elsewhere to obtain reliable information for statistical estimation.

The best practical answer seems to create a database of "simulated events". The role of simulation is to approach via a model the effects on the signal of the occurrence of different events including islanding. The simulation input is a vector of parameters (in this case mostly referring to the topology of the distribution and/or transport networks), and its output is basically the voltage and current response measured by a probe at some previously chosen nodes.

We assume the simulation model is accurate, so that the results are very close to those that a given event would create in the actual network. In this case, the quality of the database is determined by the set of parameters used.

As there is a large number of possible configurations and of possible events in a distribution network, the question of how to appropriately choose a set of parameters is not an easy one and will not be addressed in this report.

Here, we will work with an *expert constructed database*, designed to enhance the possible difficulties to correctly identify an islanding condition. We were provided a data basis of 42 simulations each accounting for 10 seconds of voltage and current signals, sampled at a rate of 10 kHz. Islanding is simulated to occur in 4 of the 42 different signals at time $t = 1s$. The remaining 38 signals correspond to other perturbations on the electrical network, different from islanding, strong enough as to induce changes in the quality of the signal.

2.2 Definition of indicators

A first visual representation of the current in both cases with and without islanding, shows that it is quite difficult to distinguish both cases from a visual inspection of the form of the signal.

This suggest that the definition of the indicators should transform the signal as to enhance possible differences when islanding occurs. On the other hand, as explained before, we need to define the indicator functions in such a way as to render them implementable in practice.

After several testing, we have focused on the spectral representation of the signal. Figure 0.1 compares a detail of the spectrogram of two signals: one that is not under islanding condition but is perturbed at time $t = 1.0s$

Indicator	Frequency range (Hz)
f_{50}	[37.5,62.5)
f_{75}	[62.5,87.5)
f_{100}	[87.5,112.5)
f_{125}	[112.5,137.5)
f_{150}	[137.5,162.5)
f_{175}	[162.5,187.5)
f_{200}	[187.5,212.5)
f_{225}	[212.5,237.5)
f_{250}	[237.5,262.5)

Table 0.1: Indicators and corresponding frequency range

and the other one for which an islanding starts at time $t = 1.0s$. Note for example that the signal after the islanding event seems to have a stronger contribution from frequencies around 150 Hz and weaker contribution below 25 Hz than either the signal before the event or the signal that has suffered a different perturbation.

Of course, this qualitative observation is not conclusive, but it helps us define our indicator functions. Let us present our proposed definition.

Each total signal is divided in signals of time length $T = 0.1024ms$ (i.e. 2^{10} samples), with an overlapping of half the window size between two consecutive windows. The data are then weighted using a Hahn function in order to focus most of the information in the central time. We recall that the Hahn function over a window of size N is given by

$$w(n) = \frac{1}{2}\left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right).$$

Then, the frequency components at each window are calculated with an FFT. Finally, frequencies are aggregated in time intervals of size 25 Hz around the harmonics and main subharmonics to form the indicators as shown in Table 0.1 . We limit our analysis to frequencies up to 250Hz. The whole process is performed in R using the *signal* package.

In addition to the previous indicators, we use the result of the traditional islanding indicators (change in frequency $f_{oldfreq}$ and change in voltage peak f_{oldv}). Hence, we will use 11 explaining variables. In addition we have the binary variable denoting the presence or absence of islanding.

Note that the whole process to calculate the indicators may be performed on-line (up to a time shift) and, the FFT being the most difficult task, can be implemented by known hardware.

2.3 Classification using the CART algorithm

Our goal is to solve a classification problem. We have chosen to work with the CART algorithm of Breiman et al. [1] (see also [3]). The algorithm is a recursive greedy algorithm that, at each iteration, looks for a binary question on one of the explanation variables that best divides the considered set into two sets that are better classified. To decide on how to split a given set, the CART algorithm uses the Gini impurity criterion: the idea is to minimize the probability of miss-classifying an element in the set given that we assign it to belong to one of the classes. Hence, the splitting is made as to lower in an important quantity the aggregated impurity after the splitting. The recursion stops either when there is no improvement (for example when all elements already belong to a group) or if the improvement is below a given threshold.

After the construction of the initial classification tree, a "pruning" stage is frequently used. The main objective of pruning is to avoid over-fitting of the classification tree with respect to the construction sample. In general, one expects a parsimonious classification tree to apply to a higher range of examples. Hence, pruning allows to change the balance between parsimony and classification error.

To end with this brief review on the CART method, we recall that it is also important to fix the relative importance of each kind of miss-classification. In our particular example, it accounts to fixing the relative importance of not detecting an existing islanding versus detecting a false islanding. If no weights are given, the relative importance is given by the data themselves. Otherwise, they are given by the relative size of each group in the sample.

The CART algorithm has several characteristics that make it well adapted to find a solution to our kind of problem. First of all the final decision tree is easily interpretable, allowing for an expert review and giving elements to understand the explained phenomenon. For example in our setup, the CART procedure suggest that islanding generates a quite strong distortion of some subharmonics. Second, using the classification tree for predictions is straightforward and quick, which is very convenient when aiming to obtain an on-line criteria: the total cost in implementation and time calculation of the islanding detector is basically that of calculating the indicators.

2.4 Numerical results

For the tree construction algorithm we apply the procedure explained for the indicator phase to the signal segment corresponding to $[0.25, 2.5]$ seconds. The interval $[0, 0.25]$ seconds is neglected to avoid the influence of simulation transient effects. The rest of the time interval is saved for posterior validation. We obtain, then, a database with 11 explaining variables, one classification variable and 6972 observations.

Let us start by considering the weighting function. Clearly, we do not have any *a priori* knowledge of how to weight both classes of errors. As we believe that islanding is a very rare effect that is over-represented in the available database (with the a proportion of about 6.4% of the available data), we conclude that we are giving already a higher importance to correctly identify islanding over giving false detections. Hence, do not add any additional weighting to the loss function of the algorithm.

We use the implementation of the CART algorithm available in the *rpart* package in R.

The application of the algorithm over the described database, proposes a tree taking into account only four indicators: f_{125}, f_{150}, f_{50} and the traditional frequency indicator $f_{oldfreq}$.

In order to evaluate if some pruning is convenient, we look at the relative and cross-validation errors at each splitting step in the CART algorithm construction. Table 0.2 shows the evaluation of the method for each splitting step. We are particularly interested in the cross-validation error, meaning the error in classification when at each construction step part of the observations are saved to evaluate the estimation made on the remaining part. We also evaluate the relative error, that is the ratio between the classification error using the tree on the whole tree over the initial rate of islanding observations.

Note that the maximal improvement in the cross-validation error is obtained at level 3. Adding additional classification levels does not reduce in a sensible way the total cross-validation error, so that we consider a good tradeoff between parsimony and precision is to use the tree at level 3.

The final tree is presented in Figure 0.2. The final classification algorithm is quite simple, and takes into account only two criteria: there will

Splits	Complexity	Rel. Error	C.V. error	C.V. SD
0	0.675000	1.000000	1.000000	0.046144
1	0.243182	0.325000	0.325000	0.026898
2	0.015909	0.081818	0.081818	0.013601
3	0.013636	0.065909	0.081818	0.013601
5	0.010000	0.038636	0.072727	0.012827

Table 0.2: Evaluation table of the CART algorithm for different depth levels

	Initial database		Validation database	
	Original	New	Original	New
False detections	7	1	0	0
Missed islandings	2	0	2	0

Table 0.3: Comparison of number of false detections and missed is landings from the 42 tested situations

be islanding if the f_{150} indicator is above a threshold $2.4dB$.

Let us test the performance of the algorithm. We perform two tests. One, with less statistical value, in which we use the same sample used for classification, and a second one, using a similar database from the observations in the interval $T = [3s, 5s]$ that were not used during the tree construction.

We define a *false detection* as the indicator signaling at least once in the analyzed time interval the presence of islanding. Likewise, we say that there is a missed islanding if the detector does not emit a detection signal for at least 10 consecutive time frames.

Finally, let us look at a graph showing each point, their group and the value of both indicators chosen for the final classification (f_{150}vs. f_{125}), as shown in Figure 0.3. Note how the classification algorithm chooses the quadrant containing most of the islanding data. On the other hand, the signals where there is no islanding is seem to be placed forming a straight line. It would be interesting to study if this apparent linear relation might be of use to develop alternative detection methods based on outlier detection. Unfortunately we did not have enough time to test this idea, so that we propose it as a possible perspective.

2.5 Conclusion

From this section we conclude that the procedure to propose a new passive islanding detection method seems promising. From the physical point of view, the classification tree obtained, based on the observation of the power in two frequency groups, seems reasonable as the signal delivered by distributed generators is of lower quality. The proposed criterion seems to be more efficient than the current existing one, as shown by the cross-validation test and is easily implementable. Automatic generation of simulated signals would be recommended in order to assure a better starting database and to perform further tests on the quality of the opposed criterion.

We believe further studies are possible. For example, it is interesting to study an alternative approach of outlier detection, profiting from some linear relation between power at two frequency groups when islanding is not present.

3 Appendix:

The lack of information on islanding (causes, places and time) at ERDF, remains a major inconvenience and a big obstacle to develop methods for identifying false or true islanding .

In this section, we will introduce a naive method to treat our problem. The main goal of this approach is rather to develop a code to perform a record for the different types of islanding (true or false).

We talked about a naive method because we made assumptions that are not real. Indeed, we consider that we already have a database where there is a real islanding with all its causes.

For each case, we adopted a specific sensor. the characteristic of sensor that will send us a logical symbol $(0 or 1)$, 1 if the cause appear and 0 if not (Figure 0.4).

In the ERDF control center, we will install a machine that will receive all the signals sent by the sensors. once that's done, we use our developed Matlab code, the purpose of this Code is to gather all the values obtained in the form of a vector V which we will calculate the norm $||V||$. According to this calculation, we will decide the nature of islanding (Figure 0.5).

- if $||V|| = 0$, then no type of causes appear and in this case the islanding detected is real (true islanding) .

- if $||V|| \neq 0$, then some types of causes appear and in this case the islanding detected is not real(false islanding).

Every time we use the code, it will make a record in a text file with the details of islanding.

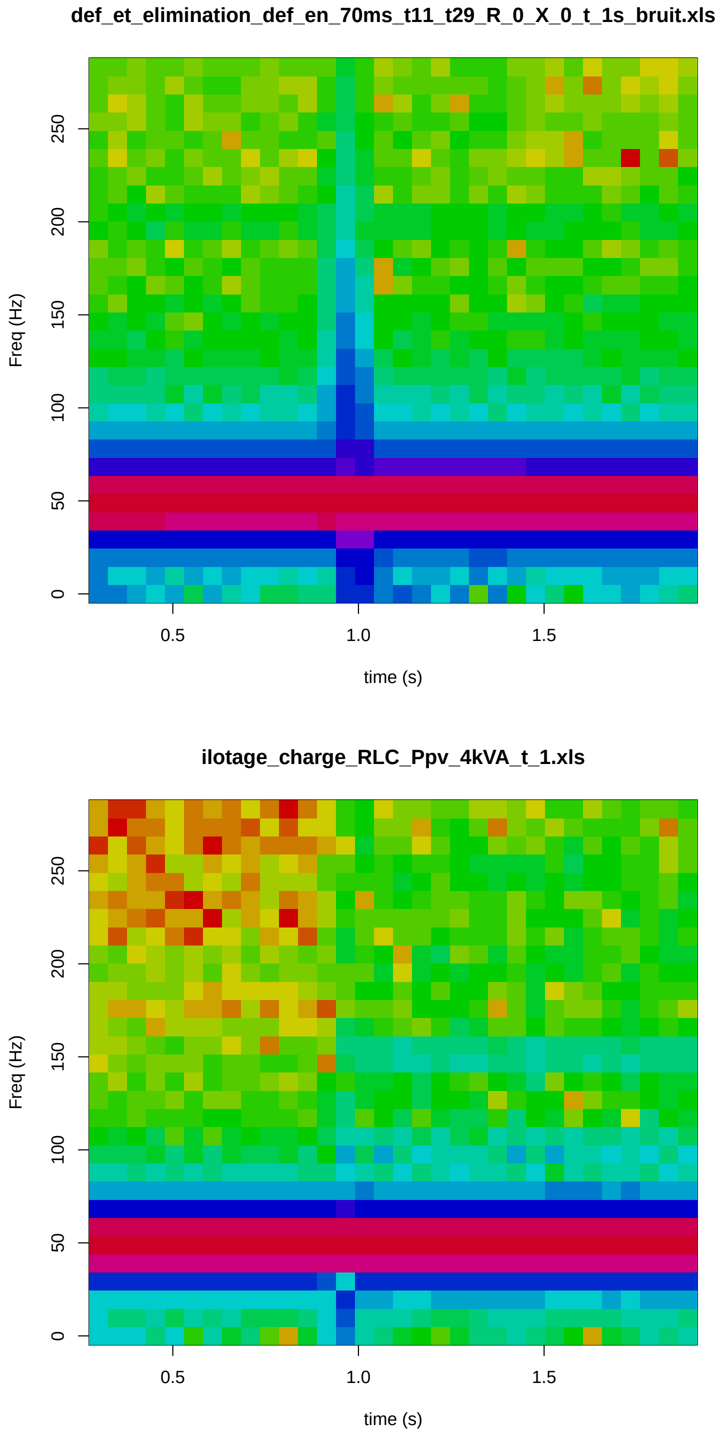


Figure 0.1: Spectrogram of two signals: up, signal without islanding; down, signal with islanding.

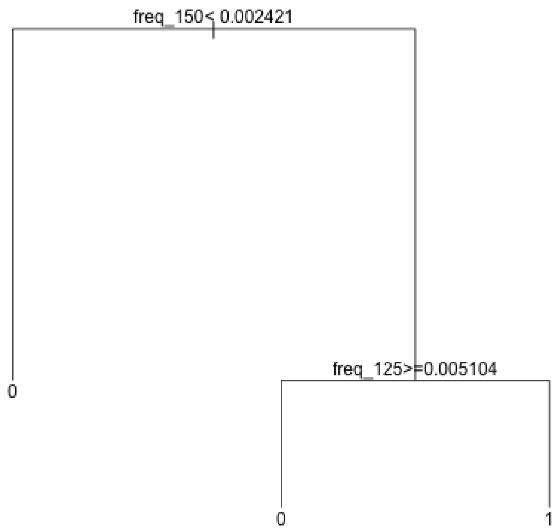


Figure 0.2: Final classification tree

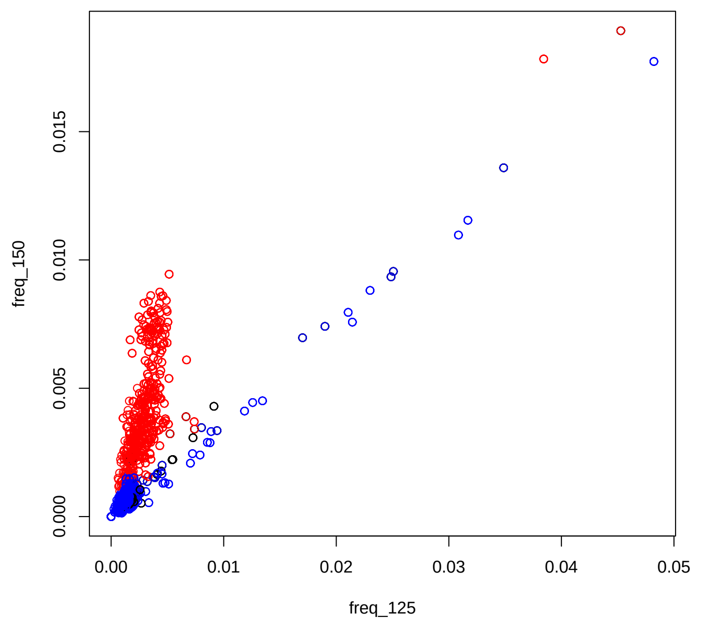


Figure 0.3: Classification variable: Red: islanding. Blue: not islanding.

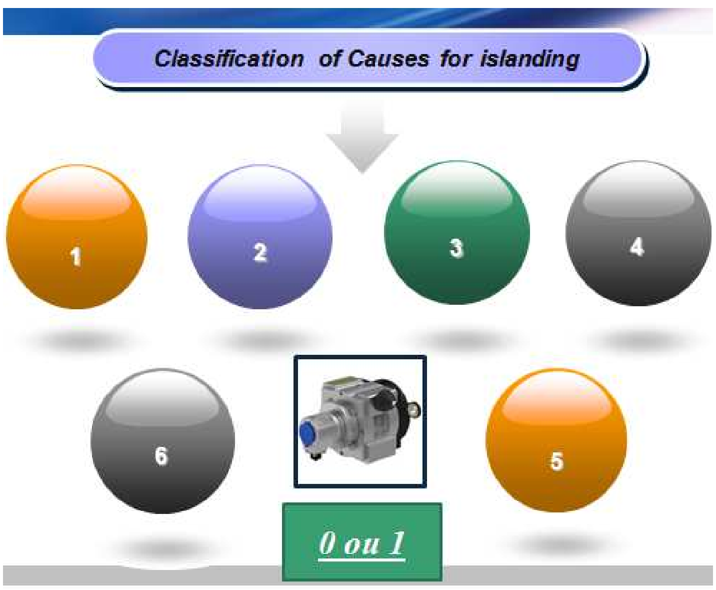
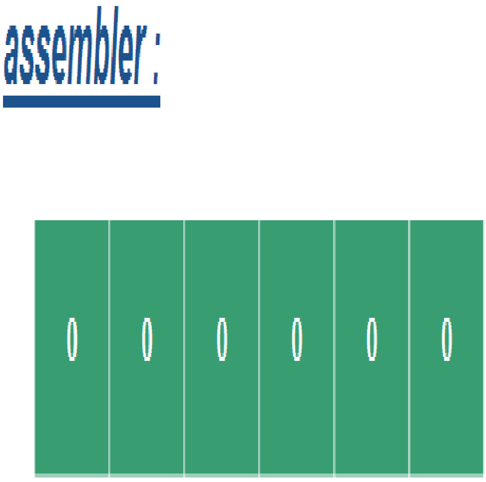
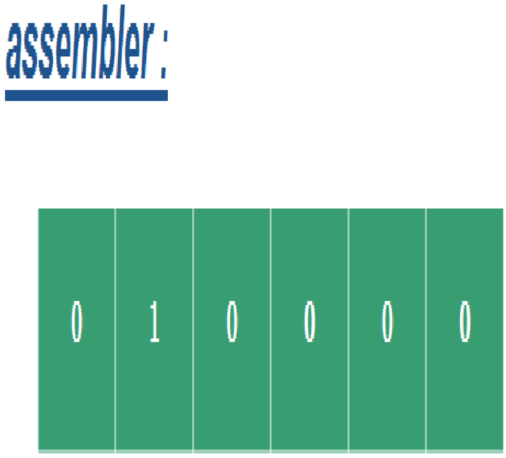


Figure 0.4: The sensor characteristic.

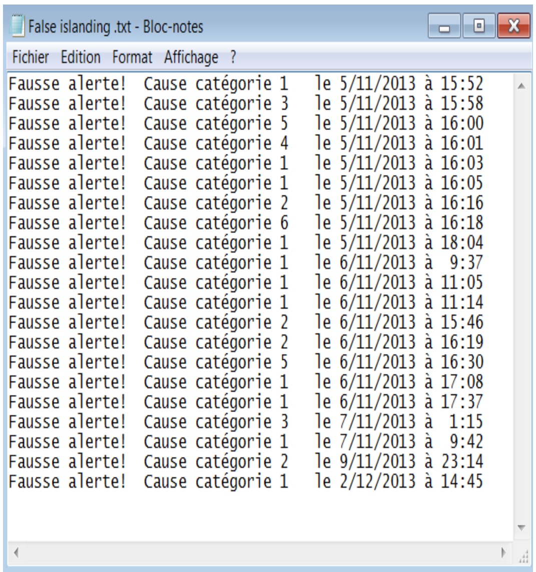


(a) False islanding



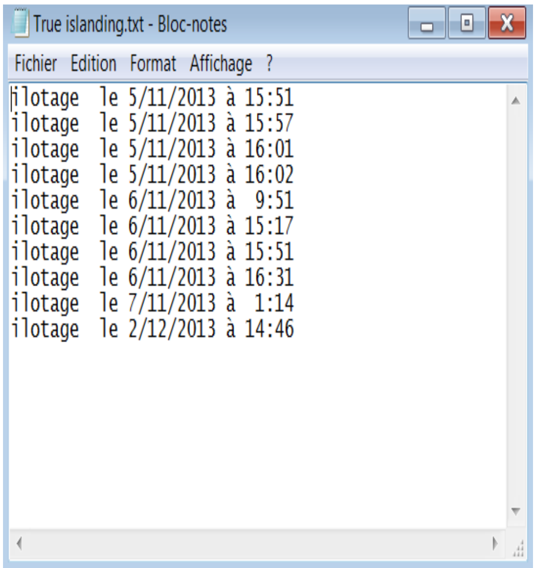
(b) True islanding

Figure 0.5: Type of islanding




```
Fausse alerte!  Cause catégorie 1   le 5/11/2013 à 15:52
Fausse alerte!  Cause catégorie 3   le 5/11/2013 à 15:58
Fausse alerte!  Cause catégorie 5   le 5/11/2013 à 16:00
Fausse alerte!  Cause catégorie 4   le 5/11/2013 à 16:01
Fausse alerte!  Cause catégorie 1   le 5/11/2013 à 16:03
Fausse alerte!  Cause catégorie 1   le 5/11/2013 à 16:05
Fausse alerte!  Cause catégorie 2   le 5/11/2013 à 16:16
Fausse alerte!  Cause catégorie 6   le 5/11/2013 à 16:18
Fausse alerte!  Cause catégorie 1   le 5/11/2013 à 18:04
Fausse alerte!  Cause catégorie 1   le 6/11/2013 à  9:37
Fausse alerte!  Cause catégorie 1   le 6/11/2013 à 11:05
Fausse alerte!  Cause catégorie 1   le 6/11/2013 à 11:14
Fausse alerte!  Cause catégorie 2   le 6/11/2013 à 15:46
Fausse alerte!  Cause catégorie 2   le 6/11/2013 à 16:19
Fausse alerte!  Cause catégorie 5   le 6/11/2013 à 16:30
Fausse alerte!  Cause catégorie 1   le 6/11/2013 à 17:08
Fausse alerte!  Cause catégorie 1   le 6/11/2013 à 17:37
Fausse alerte!  Cause catégorie 3   le 7/11/2013 à  1:15
Fausse alerte!  Cause catégorie 1   le 7/11/2013 à  9:42
Fausse alerte!  Cause catégorie 2   le 9/11/2013 à 23:14
Fausse alerte!  Cause catégorie 1   le 2/12/2013 à 14:45
```

(a) False islanding




```
ilotage  le 5/11/2013 à 15:51
ilotage  le 5/11/2013 à 15:57
ilotage  le 5/11/2013 à 16:01
ilotage  le 5/11/2013 à 16:02
ilotage  le 6/11/2013 à  9:51
ilotage  le 6/11/2013 à 15:17
ilotage  le 6/11/2013 à 15:51
ilotage  le 6/11/2013 à 16:31
ilotage  le 7/11/2013 à  1:14
ilotage  le 2/12/2013 à 14:46
```

(b) True islanding

Figure 0.6: Type of islanding - Matlab recording -

Bibliography

[1] Breiman, Leo and Friedman, Jerome H. and Olshen, Richard A. and Stone, Charles J., 1984. Wadsworth Statistics/Probability Series. Classification and regression trees. *Wadsworth Advanced Books and Software.*

[2] Therneau, Terry M. and Atkinson, Elizabeth J., 1997. An introduction to recursive partitioning using the RPART routines *Technical Report 61, Section of Biostatistics, Mayo Clinic, Rochester.*

[3] Berk, Richard A., 2008. Statistical learning from a regression perspective. *Springer-New York*. Springer Series in Statistics.